ADVANCES IN MATHEMATICS 65, 171–203 (1987)

Automorphisms of $\det(X_{ij})$: The Group Scheme Approach*

WILLIAM C. WATERHOUSE

Department of Mathematics, Pennsylvania State University, University Park, Pennsylvania 16802

Let A be a commutative ring with unit. For each $n \geqslant 2$ we can consider the form $\det(X_{ij})$, a homogeneous polynomial of degree n in the n^2 variables X_{ij}. We can then ask for the automorphisms of the determinant, the linear changes of variable that preserve the form. Clearly for instance we leave the determinant unchanged if we replace (X_{ij}) by $P(X_{ij})\,Q$ where P and Q are $n \times n$ matrices over A with $\det(PQ)=1$. The transposition mapping $(X_{ij}) \mapsto (X_{ij})^{\mathrm{tr}} = (X_{ji})$ also preserves the determinant. When A is a field, these maps give the whole automorphism group; this fact was first discovered by Frobenius [6] and has been proved afresh several times since then [4, 17, 20]. Recently James [9, 10] extended the study of this automorphism group to integral domains A, first for $n=2$ (where the determinant is a quadratic form) and then for arbitrary n; he found that there were other automorphisms corresponding to certain elements in the Picard group of A. In this paper we will extend his results and find the group for all commutative rings A. In the process we will also determine the changes of variable preserving the equations $\Lambda^r X = 0$. We will then go on to settle similar questions for symmetric and skew matrices.

This is also a programmatic paper, and the results on determinants are only part of its *raison d'être*. The fact is that automorphism (or "preserver") questions like these have been the object of quite a lot of work in linear algebra; to see this, the reader need only look at the surveys by Marcus [13, 14], or at recent papers like [23, 22, 15, 32]. The problems thus studied over fields almost always make sense over more general rings, and I hope to show that it is both reasonable and desirable to study them in that generality.

It is of course always desirable to answer a question in its natural generality, and such answers can be reached by various methods. In fact McDonald has independently determined the automorphisms of $\det(X_{ij})$

* This work was partially supported by the National Science Foundation, Grants MCS 79–03067 and MCS 8102697.

0001-8708/87 $7.50

Copyright © 1987 by Academic Press, Inc.
All rights of reproduction in any form reserved.

over general A by an ingenious and painstaking extension of one of the known proofs over fields [19]. But I want to emphasize how reasonable it is to consider such a problem over general A in the setting of group schemes. Indeed, any problem of this type is by its very definition the study of an affine group scheme. Thus it naturally fits into a framework where a certain amount is known in general, and this knowledge can be quite useful in several different ways.

First, there are general properties of affine group schemes that will necessarily turn up in such problems, and there is no point in rediscovering them anew each time. In our specific $\det(X_{ij})$ problem, for instance, the answer when properly formulated *is exactly the same* over arbitrary rings as over fields. The extra automorphisms discovered by James are automatically present because they are inherent in the group scheme involved, and they have in fact been observed in similar problems before [24, 28].

Second, a limited amount of knowledge may suffice to determine an affine group scheme, so that we may not need to consider truly arbitrary commutative rings. The reader will see, for instance, that in this paper we never need to pay any serious attention to rings with nontrivial idempotents. Indeed, most of our specific arguments will be carried out over algebraically closed fields.

A third point, more technical but often important, is that descent theory in characteristic p requires information about the automorphism group scheme beyond merely knowing the automorphisms over fields. The results in this paper, for instance, imply that every twisted form of the determinant has a separable splitting field, and this information is crucial for the determination of such twisted forms [31].

Finally, the flexibility provided by the machinery of group schemes can actually simplify the arguments needed even over fields. The theorems of this paper, for instance, obviously must contain a proof of Frobenius' original result; and this proof is quite different from those previously known. The basic difference is that we are able to reduce the problem to a linear one by passing to the Lie algebra. Such reduction is of course a familiar classical idea over the reals; affine group schemes allow us to apply the same idea in all characteristics.

In view of these advantages, I want to make it easy for others to use the methods of this paper in their own work. To this end I have concentrated all the general group scheme arguments in the first section, where they are woven into a brief exposition of the properties of affine group schemes. Some of these properties are formally proved here as theorems, but I do not claim these theorems as basically new results; they just do not happen to be stated explicitly in the source books [2, 3, 27]. The most important such result is Theorem 1.6.1, which is specifically designed for our applications. The viewpoint adopted throughout is the same as in my textbook [27], but no particular familiarity with that book is presupposed.

1. Affine Group Schemes

1.1. An affine group scheme G, for the purposes of this paper, will be a subgroup of the $n \times n$ invertible matrices $GL(n)$ defined by polynomial equations on the matrix entries. That is, we have a fixed base ring k and a fixed collection of polynomials in $k[(Y_{ij})]$, and for every k-algebra A we have a matrix group $G(A)$ consisting of the elements (y_{ij}) in $GL(n, A)$ for which the polynomials are zero. Such a G is in particular functorial; that is, an algebra map $A \to B$ induces a homomorphism $G(A) \to G(B)$. In most of our applications the defining polynomials will have integer coefficients, so that we can take $k = \mathbb{Z}$ and have G defined for all commutative rings A. More is known about group schemes defined over base fields, however, so we may still occasionally restrict to algebras over some field L; such a restriction will be denoted G_L.

1.2. The simplest example of an affine group scheme is of course $GL(n)$ itself. (The special case $n = 1$, the multiplicative group, occurs often enough that it has the special symbol G_m.) Orthogonal groups are another example. More generally, suppose f is any polynomial whatever in $k[X_1, ..., X_n]$. We can ask then for the automorphisms of f, the linear changes of variable $TX_j = \sum a_{ij} X_i$ that preserve the polynomial (i.e., satisfy $f(TX) = f(X)$). This question makes sense for a_{ij} in any k-algebra A, so at least it defines subsets $\mathrm{Aut}(f)(A)$ inside each $GL(n, A)$. Since the changes of variable give an action of the group $GL(n, A)$ on polynomials with coefficients in A, the stabilizer $\mathrm{Aut}(f)(A)$ of f is a subgroup of $GL(n, A)$. Moreover, if we write out $f(TX)$, we get a polynomial where the coefficients of the various X^α are themselves polynomial expressions involving the (fixed) coefficients of f and the matrix entries (a_{ij}) of T. The condition $f \circ T = f$ thus is equivalent to certain polynomial equations on the a_{ij} with coefficients in k. Hence $\mathrm{Aut}(f)$ is an affine group scheme defined over k.

1.3. One cannot get an affine group scheme to come out equal to a fixed (nontrivial) finite group for all A. To see why not, suppose we try to write down a group G of order 2, say the group consisting of the 2×2 matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

These two elements satisfy the polynomial equations

$$a_{11} = a_{22}, \qquad a_{11}a_{12} = 0$$

$$a_{12} = a_{21}, \qquad a_{11} + a_{12} = 1.$$

Over a field or an integral domain, our two matrices are indeed the only two satisfying the equations. But if a nontrivial idempotent e is present in A, then setting $a_{11} = e$ and $a_{12} = 1 - e$ gives another solution. In fact, the group $G(A)$ consists precisely of all matrices of the form

$$e\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + (1-e)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Such behavior is automatic, so much so that a "constant" group *scheme* of order 2 is just defined to have one element for each solution of $e^2 = e$ in A. By abuse of notation, this group scheme is often still written as $\mathbb{Z}/2\mathbb{Z}$.

For an immediately relevant example, we can consider transposition, an operation of order 2 on matrices which preserves the determinant. Since $\mathrm{Aut}(\det(X_{ij}))$ is an affine group scheme, this operation must bring others with it whenever idempotents appear in A. Once this tells us what to look for, it is then trivial to observe that the maps

$$X_{ij} \mapsto eX_{ij} + (1-e)\,X_{ji}$$

for $e^2 = e$ do indeed preserve the determinant. They give us a copy of the constant group scheme $\mathbb{Z}/2\mathbb{Z}$ sitting inside $\mathrm{Aut}(\det(X_{ij}))$.

1.4. Sometimes, rather than preserving a polynomial, we may want to preserve the condition of its being zero. More generally we may have several polynomials $f_1, \ldots, f_s$ (for instance, the various $k \times k$ minor determinants of (X_{ij})) and want to preserve the condition that says they are simultaneously zero. There are different ways to interpret this (see the Appendix), but we adopt here the algebraist's view that polynomials are formal expressions. Thus we are looking for the changes of variable T such that each $f_i(TX)$ is a sum of (polynomial) multiples of the $f_j(X)$. We may then equivalently pass to the ideal I in $A[X_1, \ldots, X_n]$ generated by $f_1, \ldots, f_s$; we say that T preserves I if $f \circ T$ is in I for every f in I.

PROPOSITION 1.4.1. *If T preserves an ideal I, so does T^{-1}.*

Proof. It is known [21] that an A-algebra endomorphism of a finitely generated A-algebra is an isomorphism whenever it is surjective. Now if composition with T preserves I, the isomorphism it gives on $A[X_1, \ldots, X_n]$ passes to an endomorphism of $A[X_1, \ldots, X_n]/I$. This endomorphism is clearly surjective, and hence it is bijective. This implies that T^{-1} preserves I. ∎

A special form of this result was proved by Dixon [5].

COROLLARY 1.4.2. *Suppose $f_1,\dots,f_s$ in $A[X_1,\dots,X_n]$ are all homogeneous of the same degree. Then any linear change of variables preserving the ideal $I=(f_1,\dots,f_s)$ induces a bijective linear map on the A-span of the f_i.*

Proof. The elements $\sum c_i f_i$ with c_i in A are the only elements in I of the same degree as the f_i. Hence the linear change of variables maps the set of these elements to itself. The inverse change also preserves I; thus it also maps this set to itself, and the two maps are inverse to each other there. ∎

Suppose now that we start with $f_1,\dots,f_s$ in $k[X_1,\dots,X_n]$. For every k-algebra A we can then consider the subset $G(A)$ in $GL(n,A)$ preserving the ideal generated by the f_i. Lemma 1.4.1 shows that each $G(A)$ is a group. Moreover, if $A\to B$ is a k-algebra map, it clearly sends $G(A)$ into $G(B)$. In general, however, the $G(A)$ may fail to form an affine group scheme. The reason for this is that the preservation of an ideal may impose divisibility conditions, as well as polynomial equations, on the matrix entries. For example, take the polynomials X_1^2, X_2^2, $4X_1X_2$ over the integers. It is easy to compute that the ideal these generate is preserved precisely when $2a_{11}a_{21}$ and $2a_{12}a_{22}$ are divisible by 4. Over the rationals, all (a_{ij}) satisfy this condition, and thus no nontrivial polynomial over the integers can vanish on all the solutions. Hence $G(A)$ is not definable by polynomial equations. This example shows that we need here some technique for proving when the groups preserving an ideal form an affine group scheme. The following simple test is usually enough.

THEOREM 1.4.3. *Let V be the module of all polynomials in $k[X_1,\dots,X_n]$ of degree at most r. Let I be an ideal generated by elements of degree at most r. Suppose there is a basis $f_1,\dots,f_t$ of V such that $f_1,\dots,f_s$ among them span $I\cap V$. Then the condition of preserving I defines an affine group scheme on k-algebras.*

Proof. Suppose T preserves I. Take the elements $f_i\circ T$ for $i\leqslant s$, which are in V; when we expand them out in the basis $f_1,\dots,f_t$, they must have 0 as the coefficient of f_j for each $j>s$. Conversely, this zero-coefficient condition forces T to map $I\cap V$ into itself and hence preserve I. But clearly the zero-coefficient condition is just a collection of polynomial equations on the coefficients of T. ∎

COROLLARY 1.4.1. *If k is a field, the condition of preserving an ideal in $k[X_1,\dots,X_n]$ always defines an affine group scheme on k-algebras.*

Porism 1.4.5. *If I is generated by homogeneous polynomials of degree r, then the theorem remains true if V is replaced by the space of all polynomials homogeneous of degree r.*

Corollary 1.4.6. *Suppose in the theorem we have $k=\mathbb{Z}$. Suppose that the rank of $I \cap V$ is the same when we read the polynomials modulo each prime. Then the condition of preserving I defines an affine group scheme. This is true in particular if I is generated by homogeneous polynomials of degree r that remain independent modulo p for all p.*

Proof. The structure theory for subgroups of $\mathbb{Z}^t$ shows us that there is a basis $f_1, \ldots f_t$ of V together with nonzero $b_1, \ldots, b_s$ in $\mathbb{Z}$ such that $b_1 f_1, \ldots, b_s f_s$ is a basis of $I \cap V$. The polynomials in $I \cap V$ then have rank lower than s modulo some prime p iff p divides some b_i. We have assumed that this never happens, and hence all b_i are ± 1 and the hypothesis of the theorem is satisfied. ∎

For a relevant example, now, suppose we take mn indeterminates X_{ij} with $i=1, \ldots, m$ and $j=1, \ldots, n$. For each r between 1 and $\min(m, n)$, let $H_r(A)$ denote the invertible linear maps over A that preserve the ideal generated by all $r \times r$ minor determinants of (X_{ij}). These different minors are independent modulo every p, since in fact we can assign values to the X_{ij} to make any prescribed minor 1 while all the others vanish. Hence Corollary 1.4.6 can be applied:

Corollary 1.4.7. *The groups $H_r(A)$ are an affine group scheme.*

Even when the hypothesis of (1.4.6) is not satisfied, we are never far from affine group schemes. To see this, observe that the proof of (1.4.6) remains valid when k is any principal ideal domain; in particular, the result is true over localizations of $\mathbb{Z}$. Now if we start with an ideal over $\mathbb{Z}$ and construct the basis $b_i f_i$ of $I \cap V$, we can choose an integer q containing all prime factors of all the b_i. Then over $\mathbb{Z}[1/q]$ we can apply the result:

Corollary 1.4.8. *Let I be an ideal of $\mathbb{Z}[X_1, \ldots, X_n]$. Then there is some integer $q>0$ such that the invertible linear maps preserving I are an affine group scheme on all $\mathbb{Z}[1/q]$-algebras.*

1.5. Suppose in this subsection that our base ring k is a field. Let G be an affine group scheme over k. Form the ring $k[\varepsilon]$ with $\varepsilon^2=0$, and look at the elements in $G(k[\varepsilon])$ that reduce to the identity in $G(k)$ when we send ε to 0. These form a k-space that is called the Lie algebra of G and denoted $\operatorname{Lie}(G)$. (It does in fact carry a Lie algebra structure, but that is irrelevant for our purposes.) In particular, $\operatorname{Lie}(GL(n))$ consists of all matrices of the form $T = I + \varepsilon M$ where $M = (m_{ij})$ is a matrix in k. For our subgroup G of $GL(n)$, we can then find $\mathrm{Lie}(G)$ by computing which $I + \varepsilon M$ satisfy the condition defining G. The fact that $\varepsilon^2 = 0$ usually makes this computation easy. For instance, suppose we take G to be $\mathrm{Aut}(f)$, where $f(X_1, ..., X_n)$ is some polynomial. It is trivial to check that

$$f((I + \varepsilon M)\, X) = f(X) + \varepsilon \sum_j \left(\frac{\partial f}{\partial X_j}\right)\left(\sum_i m_{ij} X_i\right).$$

Thus $\mathrm{Lie}(\mathrm{Aut}(f))$ is given by the (m_{ij}) satisfying

$$\sum_{ij} m_{ij} X_i \left(\frac{\partial f}{\partial X_j}\right) = 0.$$

The dimension $\dim(G)$ of an affine group scheme G is defined in essence to be the number of independent variables necessary to specify an element of G; there are several equivalent ways of making this precise. The vector space $\mathrm{Lie}(G)$ has linear dimension at least as large as $\dim(G)$, and one says G is smooth if $\dim \mathrm{Lie}(G) = \dim(G)$. This condition is automatic when $\mathrm{char}(k) = 0$, and the smooth groups escape many of the peculiarities that can otherwise occur in positive characteristic.

Inside G one can define a normal subgroup G^0 called the connected component (of the identity). One says G is connected if $G = G^0$; the group $GL(n)$, for instance, is connected. In any case G^0 has finite index, and it has the same Lie algebra as G. Any homomorphism from a connected group scheme F to G will factor through G^0.

1.6. We can now prove the specific isomorphism result that will be used in this paper. It contains a requirement that our group schemes be "of finite type," but that is included just to assuage the experts; it is automatically true for the subgroups of $GL(n)$ that we have been considering. There is also one assumption of flatness; this will be automatic in our cases, where G will be G^0 or $G^0 \rtimes (\mathbb{Z}/2\mathbb{Z})$ and the G^0_L will all be smooth and connected of the same dimension [3, Vol. I, p. 349].

THEOREM 1.6.1. *Let G and H be affine group schemes of finite type over $\mathbb{Z}$ with G flat, and let $\varphi: G \to H$ be a homomorphism. Suppose that for all algebraically closed fields L we can show that*

(i) $\dim(G_L) \geqslant \dim_L \mathrm{Lie}(H_L)$,

(ii) *the maps $G(L) \to H(L)$ and $G(L[\varepsilon]) \to H(L[\varepsilon])$ given by φ are injective (where $\varepsilon^2 = 0$), and*

(iii) *all elements inside $H(L)$ normalizing $\varphi(G^0(L))$ are in $\varphi(G(L))$.*

Then φ is an isomorphism; that is, it sends $G(A)$ isomorphically onto $H(A)$ for every commutative ring A.

Proof. The crucial idea in this proof is that (by an equivalent form of the definition) every affine group scheme G over $\mathbb{Z}$ has associated with it a commutative ring $\mathbb{Z}[G]$ with $G(A) = \mathrm{Hom}_{\mathrm{rings}}(\mathbb{Z}[G], A)$. Thus we can prove isomorphism by proving that the ring map $\Phi\colon \mathbb{Z}[H] \to \mathbb{Z}[G]$ induced by φ is an isomorphism.

We must begin by proving isomorphism over fields, and for this several results in [27] will be used without specific citation. First, (ii) implies $\dim(G_L) \leqslant \dim(H_L)$. Together with (i), this implies that G_L and H_L are both smooth and have the same dimension. By (ii) then $\varphi(G^0(L)) = H^0(L)$. But $H(L)$ normalizes $H^0(L)$, so by (iii) we have $\varphi(G(L)) = H(L)$. As G_L and H_L are smooth, they are determined by their points in L, and (ii) for $L[\varepsilon]$ makes the map between them separable; hence $G_L \to H_L$ is an isomorphism. Then if k is any field, descent from $L = \bar{k}$ to k shows that $G_k \to H_k$ is an isomorphism.

Flatness of G now means exactly that $\mathbb{Z}[G]$ is a torsion-free $\mathbb{Z}$-module. The map $\mathbb{Z}[H] \to \mathbb{Z}[G]$ we know becomes an isomorphism when tensored with the field $\mathbb{Q}$, so its kernel M consists precisely of the torsion elements in $\mathbb{Z}[H]$. As $\mathbb{Z}[H]$ is a finitely generated $\mathbb{Z}$-algebra, the ideal M is a finitely generated $\mathbb{Z}[H]$-module, and hence there is a bound on the orders of elements in M. Suppose for some prime p there is a nontrivial p-primary component of M. Then we can choose an element of largest p-power order, and it will give a nontrivial class in M/pM. But clearly $p\mathbb{Z}[H] \cap M = pM$, and so we get a nontrivial element in the kernel of $\mathbb{Z}[H]/p\mathbb{Z}[H] \to \mathbb{Z}[G]/p\mathbb{Z}[G]$. This is impossible, because we know the map becomes an isomorphism over the field $\mathbb{Z}/p\mathbb{Z}$. Hence M is zero and $\Phi\colon \mathbb{Z}[H] \to \mathbb{Z}[G]$ is injective. To show the injection is onto, suppose we could find some x in $\mathbb{Z}[G]$ not in the image. We know the map becomes an isomorphism over $\mathbb{Q}$, so by appropriate multiplication we can assume some px is in the image, say $px = \Phi(y)$. But px becomes zero when we tensor with $\mathbb{Z}/p\mathbb{Z}$. So the same must be true of y. Hence $y = pz$ for some z, and x must equal $\Phi(z)$ since $\mathbb{Z}[G]$ is torsion-free. ∎

The proof actually requires hypotheses (i), (ii), and (iii) only for the algebraic closures of the prime fields, but generally it is no harder to prove them for all algebraically closed L. It may be worth noting, though, that the same argument will work for G and H defined over $\mathbb{Z}[1/q]$ if we verify the hypotheses for those L whose characteristics do not divide q.

1.7. Some of our group schemes will be constructed as quotients of others, and one warning is needed about this construction. We start with N, a normal subgroup scheme of G, so that $N(A)$ is a normal subgroup of

$G(A)$ for each A. In our cases (and in any case when the base ring is a field) a quotient G/N can be defined with the appropriate universal property; any group scheme homomorphism $G \to H$ with kernel N will give an injection of G/N into H. Over a base field also $\dim(G/N)$ will equal $\dim(G) - \dim(N)$. But the group $(G/N)(A)$ will in general be bigger than $G(A)/N(A)$. Formally, this is just like constructing a quotient sheaf rather than a presheaf quotient. Explicitly, it means that some elements of $(G/N)(A)$ may come not from $G(A)$ but from some larger $G(B)$ when $A \to B$ is a ring extension.

In general one can say only that the $A \to B$ involved can be taken faithfully flat. But in each particular case we can be more specific. Certain cohomology sets $H^1(A, G)$ can be defined that fit into an exact sequence

$$1 \to N(A) \to G(A) \to (G/N)(A) \to H^1(A, N) \to H^1(A, G) \to H^1(A, G/N).$$

Thus the failure of $G(A) \to (G/N)(A)$ to be surjective is measured by the elements in $H^1(A, N)$ that become trivial in $H^1(A, G)$. An element in $(G/N)(A)$ will come from $G(B)$ if the corresponding class in $H^1(A, N)$ becomes trivial in $H^1(B, N)$.

The sets $H^1(A, G)$ are important for their own sake in descent theory: if G is the automorphism group of some structure, then $H^1(A, G)$ classifies the twisted forms of that structure defined over A. The most important case for us here is $G = G_m$, which is the automorphism group of the free module of rank 1. In this case $H^1(A, G_m)$ is $\mathrm{Pic}(A)$, the Picard group of projective rank 1 A-modules.

2. Some Normalizer Computations

To apply Theorem 1.6.1 we need information on normalizers, and the three results in this section are the three needed in this paper. They are not tied to our specific questions, however, and they should be of use in many similar problems.

THEOREM 2.1. *Let L be a infinite field. Let $G^0(L)$ be the collection of maps on the mn-dimensional space of $m \times n$ matrices X given by $X \mapsto PXQ^{\mathrm{tr}}$ for P in $GL(m, L)$ and Q in $GL(n, L)$. If $m \neq n$, then $G^0(L)$ is its own normalizer in $GL(mn, L)$. If $m = n$, then $G^0(L)$ is of index 2 in its normalizer, the other coset being represented by $T(X) = X^{\mathrm{tr}}$.*

Proof. We may assume $m, n > 1$. Our first step is to show that the centralizer of $G^0(L)$ consists of scalars. For this we look at the action of the subgroup H given by all diagonal elements $P = \mathrm{diag}(a_1, \dots, a_m)$ and $Q = \mathrm{diag}(b_1, \dots, b_n)$. On the basis matrices X_{ij} we have $PX_{ij}Q^{\mathrm{tr}} = a_i b_j X_{ij}$. Thus each X_{ij} spans an eigenspace for a different character of H. Any map

T commuting with $G^0(L)$ commutes with H; hence T preserves these eigenspaces, and $T(X_{ij})$ has the form $\alpha_{ij}X_{ij}$. If we now take P and Q instead to be permutation matrices, we see that α_{ij} is the same for all i and all j.

Next we recall some known results on automorphisms. As we will discuss at more length in Section 3, there is an affine group scheme G^0 for which $G^0(L)$ is the set of values in L. This G^0 is a quotient of $GL(m) \times GL(n)$, and hence it is determined on all L-algebras by its values in L (see [27, p. 115]). The automorphisms of G^0_L have been determined [29], there being only finitely many classes of them modulo conjugations by elements in $G^0(L)$. Specifically, if we let $[P, Q]$ denote the element $X \mapsto PXQ^{tr}$ then for $m \neq n$ and $m, n \neq 2$ the only nontrivial class is represented by

$$[P, Q] \mapsto [(P^{tr})^{-1}, (Q^{tr})^{-1}].$$

For $m \neq n$ and (say) $m = 2$ the only nontrivial class is represented by

$$[P, Q] \mapsto (\det P)^{-1} [P, (Q^{tr})^{-1}].$$

For $m = n$ there are three nontrivial classes; when $m \neq 2$ they are represented by

$$\begin{aligned} &[P, Q] \mapsto [Q, P], \\ &[P, Q] \mapsto [(P^{tr})^{-1}, (Q^{tr})^{-1}], \\ &[P, Q] \mapsto [(Q^{tr})^{-1}, (P^{tr})^{-1}], \end{aligned}$$

and when $m = 2$ they are represented by

$$\begin{aligned} &[P, Q] \mapsto [Q, P], \\ &[P, Q] \mapsto (\det P)^{-1} (\det Q)^{-1} [P, Q], \\ &[P, Q] \mapsto (\det P)^{-1} (\det Q)^{-1} [Q, P]. \end{aligned}$$

Now comes the point of the argument. Take any T in $GL(mn, L)$ normalizing $G^0(L)$. Conjugation by T is an algebraic automorphism of $G^0(L)$ and hence induces an automorphism φ_T of G^0_L. The centralizer computation shows that φ_T determines T up to scalars; and the scalars are in $G^0(L)$, so to find the whole normalizer we just have to determine which automorphisms occur as such φ_T. Multiplying T by an appropriate element of $G^0(L)$ will change φ_T by any desired inner automorphism, and so we just need to ask which of our representative automorphisms can occur. In the case $m \neq n$ the nontrivial class is not the identity on scalars, and thus it cannot be induced by conjugation inside $GL(mn, L)$. In the case $m = n$ the last two nontrivial classes cannot occur for the same reason, and the remaining class is realized by $T(X) = X^{tr}$. ∎

THEOREM 2.2. *Let L be an algebraically closed field. Let $G(L)$ be the collection of maps on the space of $n \times n$ symmetric matrices X given by $X \mapsto PXP^{tr}$ for P in $GL(n, L)$. Then $G(L)$ is its own normalizer in $GL(n(n+1)/2, L)$.*

Proof. We may assume $n > 1$. The proof follows the same lines as in (2.1). First, look at the subgroup of diagonal elements $P = \mathrm{diag}(a_1, ..., a_n)$. Such an element sends X_{ij} to $a_i a_j X_{ij}$, and thus the X_{ij} for $i \leqslant j$ span eigenspaces for different characters. Hence any T commuting with $G(L)$ has the form $T(X_{ij}) = \alpha_{ij} X_{ij}$. If P is the matrix for a permutation π, then $PX_{ij}P^{tr} = X_{\pi(i), \pi(j)}$. This implies that all α_{ij} with $i \neq j$ are equal, say $= \alpha$; it likewise implies that all α_{ii} are equal, say $= \beta$. Finally, we take

$$P = \left(\begin{array}{cc|c} 1 & 1 & \\ 0 & 1 & 0 \\ \hline \multicolumn{2}{c|}{0} & I \end{array}\right).$$

Then $P(X_{11} + X_{22}) P^{tr}$ is $2X_{11} + (X_{12} + X_{21}) + X_{22}$, and T sends this to $2\beta X_{11} + \alpha(X_{12} + X_{21}) + \beta X_{22}$. But

$$\begin{aligned} P(T(X_{11} + X_{22}))\, P^{tr} &= P(\beta X_{11} + \beta X_{22})\, P^{tr} \\ &= 2\beta X_{11} + \beta(X_{12} + X_{21}) + \beta X_{22}. \end{aligned}$$

As these must agree, $\alpha = \beta$, and thus our T centralizing $G(L)$ is scalar.

As in (2.1), now, we just need to examine which automorphisms of the group scheme G_L corresponding to $G(L)$ are induced by conjugations in $GL(n(n+1)/2, L)$. For $n \neq 2, 4$ we know [29] that the only nontrivial outer automorphism class is represented by the map sending the image $[P]$ of P to $[(P^{tr})^{-1}]$; this cannot occur, because it is nontrivial on scalars. Similarly for $n = 2$ the only nontrivial outer automorphism class is represented by

$$[P] \mapsto (\det P)^{-2}\, [P] = [(\det P)^{-1}\, P],$$

which is nontrivial on scalars. For $n = 4$, finally, there are three nontrivial classes represented by

$$\begin{aligned} &[P] \mapsto [(P^{tr})^{-1}], \\ &[P] \mapsto (\det P)^{-1}\, [P], \\ &[P] \mapsto (\det P)[(P^{tr})^{-1}]. \end{aligned}$$

The first two are nontrivial on scalars. Suppose the third occurs, so $T^{-1}[P]\,T = (\det P)[(P^{\mathrm{tr}})^{-1}]$. Take P to be of the form $\mathrm{diag}(a_1, a_2, a_3, a_4)$ and apply the maps to X_{11}; this shows that $T(X_{11})$ is an eigenvector for these diagonal P, with character $(a_1 a_2 a_3 a_4)\, a_1^{-2}$. This is impossible, as the X_{ij} are a basis of eigenvectors and none has this character. ∎

THEOREM 2.3. *Let L be an algebraically closed field. Let $G(L)$ be the collection of maps on the space of $n \times n$ alternating matrices X given by $X \mapsto PXP^{\mathrm{tr}}$ for P in $GL(n, L)$. If $n \neq 4$, then $G(L)$ is its own normalizer in $GL(n(n-1)/2, L)$. If $n = 4$, then $G(L)$ is of index 2 in its normalizer, the other coset being represented by*

$$T_0(X_{12}) = X_{34}, \qquad T_0(X_{13}) = -X_{24}, \qquad T_0(X_{14}) = X_{23},$$

$$T_0(X_{34}) = X_{12}, \qquad T_0(X_{24}) = -X_{13}, \qquad T_0(X_{23}) = X_{14}.$$

The proof of (2.3), very similar to that of (2.2), is contained in [29], though there the alternating matrices are identified with $\Lambda^2(L^n)$. It is interesting to note that the nontrivial automorphism for $n = 4$ here is the same one that had to be discussed separately for the symmetric case in (2.2).

3. AUTOMORPHISMS OF DETERMINANT IDEALS

Rather than studying just the automorphisms of the determinant, we will find it equally easy to treat a whole family of such problems together. We take mn indeterminates X_{ij}, which we view as the entries of a matrix X. The exterior power $\Lambda^r X$ then has entries given by the $r \times r$ minors of X. For $2 \leqslant r \leqslant \min(m, n)$, we define $H_r(A)$ to be the group of invertible linear maps on the X_{ij} over A that formally preserve the condition $\Lambda^r X = 0$. In other words, $H_r(A)$ consists of the maps preserving the ideal generated by the $r \times r$ minors. We know from (1.4.7) that each H_r is actually an affine group scheme, and our goal in this section is to determine them all. The following lemma helps indicate why we should treat them all at once.

LEMMA 3.1. *One has $H_2 \supseteq H_3 \supseteq H_4 \supseteq \cdots$.*

Proof. This proof is in essence taken from [16]. Suppose $2 < r$ and T in $H_r(A)$. Write $Y_{ij} = T(X_{ij})$. For index subsets I and J of cardinality r, let X_{IJ} denote the (I, J)-minor determinant of X. Then $\partial X_{IJ}/\partial X_{ij}$ is 0 unless $i \in I$ and $j \in J$, in which case it is $\pm X_{I-\{i\}, J-\{j\}}$. As T is in $H_r(A)$, we know by (1.4.2) that the $r \times r$ minors of $Y = (Y_{ij})$ are A-linear combinations of those of X; say $Y_{IJ} = \sum s_{IJ}^{KL} X_{KL}$. Then

$$\partial Y_{IJ}/\partial Y_{ij} = \sum s_{IJ}^{KL}(\partial X_{KL}/\partial Y_{ij})$$

$$= \sum \sum s_{IJ}^{KL}(\partial X_{KL}/\partial X_{pq})(\partial X_{pq}/\partial Y_{ij}),$$

where of course the $\partial X_{pq}/\partial Y_{ij}$ are just the coefficients of T^{-1}. The nonzero $\partial X_{KL}/\partial X_{pq}$ are (up to sign) $(r-1) \times (r-1)$ minors of X, and we can get an arbitrary $(r-1) \times (r-1)$ minor of Y in the form $\pm \partial X_{IJ}/\partial Y_{ij}$, so the $(r-1) \times (r-1)$ minors of Y are linear combinations of those for X. Hence T is in $H_{r-1}(A)$. ∎

The only real work needed for our theorem is contained in Lemma 3.2. The point to observe is that, once the first paragraph sets up the proof, we then have merely a family of linear conditions on coefficients M_{ij}^{kl}. If we wanted to establish our theorem just for some particular size of matrix, we could write out the system of equations and give it to a computer or a sophomore for solution. We have a little work to do only because we want to know about the solutions for all m and n. It can be done in various ways, and to illustrate this we use one method here and two different ones in the two corresponding later lemmas (5.3 and 6.3).

LEMMA 3.2. *Let L be a field. Then the Lie algebra of H_2 over L has dimension at most $m^2 + n^2 - 1$.*

Proof. As defined in (1.5), the Lie algebra consists of certain maps $T(X) = X + \varepsilon M(X)$ where $\varepsilon^2 = 0$. In our case $M(X)$ (like X) is a matrix $(M(X)_{ij})$ with $M(X)_{ij}$ of the form $\sum M_{ij}^{kl} X_{kl}$, and we want the T that preserve the ideal of 2×2 minors of X. Generators of that ideal have the form $X_{ik}X_{jl} - X_{il}X_{jk}$ with $i \neq j$ and $k \neq l$. Applying T to such a generator, we get

$$(X_{ik}X_{jl} - X_{il}X_{jk}) + \varepsilon(M(X)_{ik}X_{jl} + M(X)_{jl}X_{ik} - M(X)_{il}X_{jk} - M(X)_{jk}X_{il}).$$

Thus we get elements in the Lie algebra when we have values of M_{ij}^{kl} that make the coefficient of ε here lie in the ideal of 2×2 minors.

This requirement on M implies in particular that the coefficient of ε must vanish when we assign to the X_{ij} values that give a matrix of rank less than 2. For instance, suppose we set one particular X_{ik} equal to 1 and all others zero. Then for $j \neq i$ and $l \neq k$ our coefficient has just one nontrivial term, and we deduce $M_{jl}^{ik} = 0$. In other words, $M(X)_{jl}$ involves only terms in row j

and column l of the matrix X. Suppose now we fix $i \neq j$ and k, and we set $X_{ik} = X_{jk} = 1$ with all others zero. Then for each l different from k, the vanishing of the coefficient tells us that $M_{jl}^{ik} + M_{jl}^{jk} = M_{il}^{ik} + M_{il}^{jk}$. Dropping the two known to be zero, we get $M_{jl}^{jk} = M_{il}^{ik}$. Thus M_{jl}^{jk} equals M_{1l}^{1k} for all j and all $k \neq l$. A similar argument shows $M_{jl}^{il} = M_{j1}^{i1}$ for all $i \neq j$. Thus all M_{jl}^{ik} except those of the form M_{rs}^{rs} are expressible in terms of the $m(m-1)+n(n-1)$ variables M_{j1}^{i1} and M_{1l}^{1k}.

Finally, set $X_{11} = X_{1l} = X_{j1} = X_{jl} = 1$ with all other entries zero. The vanishing of the coefficient for $i = k = 1$ gives us the relation

$$\begin{aligned} 0 = {} & M_{11}^{11} + M_{11}^{1l} + M_{11}^{j1} + M_{11}^{jl} + M_{jl}^{11} + M_{jl}^{1l} + M_{jl}^{j1} + M_{jl}^{jl} \\ & - [M_{1l}^{11} + M_{1l}^{1l} + M_{1l}^{j1} + M_{1l}^{jl} + M_{j1}^{11} + M_{j1}^{1l} + M_{j1}^{j1} + M_{j1}^{jl}]. \end{aligned}$$

We already know that some of these terms are zero or cancel in pairs, and we are left then with

$$0 = M_{11}^{11} + M_{jl}^{jl} + M_{1l}^{1l} + M_{j1}^{j1}.$$

Hence coefficients of the form M_{jl}^{jl} can be expressed in terms of the $1+(m-1)+(n-1)$ variables M_{11}^{11}, M_{1l}^{1l}, and M_{j1}^{j1}. All in all, then, our requirements on M leave at most $m(m-1)+n(n-1)+1+(m-1)+(n-1) = m^2+n^2-1$ variables free. ∎

We can now introduce our candidate for the H_r and prove that it is right. Inside $GL(m) \times GL(n)$ there is a copy of G_m consisting of pairs $(\lambda I, \lambda^{-1} I)$ for invertible scalar λ, and by [3, Vol. 2, p. 15] there exists a group scheme quotient G^0 for that normal subgroup. If $m = n$, then also $\mathbb{Z}/2\mathbb{Z}$ acts on $GL(m) \times GL(n)$ by interchanging factors; this action passes to an action on G^0 and allows us to define a semidirect product.

Definition 3.3. *For* $m \neq n$, *set* $G = G^0 = (GL(m) \times GL(n))/G_m$. *For* $m = n$, *set* $G = G^0 \rtimes \mathbb{Z}/2\mathbb{Z}$.

Theorem 3.4. *For every r one has* $G \xrightarrow{\sim} H_r$.

Proof. For P in $GL(m, A)$ and Q in $GL(n, A)$ we get a transformation $X \mapsto PXQ^{\mathrm{tr}}$, and thus we have a homomorphism $GL(m) \times GL(n) \to H_r$. It is easy to check that its kernel is $(\lambda I, \lambda^{-1} I)$, and so it factors to give us an injective homomorphism $G^0 \to H_r$. When $m = n$, the map $X \mapsto X^{\mathrm{tr}}$ is also in H_r; as we noted in (1.3), this is actually a copy of $\mathbb{Z}/2\mathbb{Z}$ sitting inside H_r. We have then a homomorphism

$$[GL(m) \times GL(n)] \rtimes \mathbb{Z}/2\mathbb{Z} \to H_r,$$

and it passes to a homomorphism $G \to H_r$.

Because of the factor $\mathbb{Z}/2\mathbb{Z}(A)$, it is not at once evident that $G(A) \to H_r(A)$ is injective when $m = n$. If L is an algebraically closed field, though, then $GL(m, L) \times GL(n, L)$ maps onto $G^0(L)$, and it is simple to check in this case that $X \mapsto X^{tr}$ is not in $G^0(L)$. Hence $G(L) \mapsto H(L)$ is injective. The same is true for $G(L[\varepsilon]) \to H_r(L[\varepsilon])$. The construction of G^0 shows that $\dim G_L = m^2 + n^2 - 1$. By Lemma 3.2, $\dim G_L \geqslant \dim_L \mathrm{Lie}(H_2)_L$, and by Lemma 3.1 the same inequality holds for all higher H_r. Finally, Theorem 2.1 shows that the normalizer of $G^0(L)$ inside all of $GL(mn, L)$ is $G(L)$. Thus all the hypotheses needed for Theorem 1.6.1 are satisfied. ∎

The simple statement of Theorem 3.4 is possible only because of the group scheme properties that are included in the definition of G, and it will be worth while to make them explicit. First, when $m = n$, the group $(\mathbb{Z}/2\mathbb{Z})(A)$ in general has more than two elements; as we observed back in (1.3), it contains all the operations of the form

$$X_{ij} \mapsto eX_{ij} + (1 - e) X_{ji}$$

for $e^2 = e$. Observe, however, that no special checking is required to show that these elements have no overlap with $G^0(A)$, for that is part of what the theorem asserts.

More important is the fact that G^0 is defined as a *group scheme* quotient, and hence

$$GL(m, A) \times GL(n, A) \to G^0(A)$$

may not always be surjective. In fact, as we said in (1.7), the exact sequence

$$1 \to G_m \to GL(m) \times GL(n) \to G^0 \to 1$$

gives us

$$\begin{aligned} 1 \to G_m(A) \to GL(m, A) \times GL(n, A) \to G^0(A) \to H^1(A, G_m) \\ \to H^1(A, GL(m) \times GL(n)). \end{aligned}$$

We also observed there that $H^1(A, G_m)$ is the Picard group $\mathrm{Pic}(A)$. We have

$$H_1(A, GL(m) \times GL(n)) = H^1(A, GL(m)) \times H^1(A, GL(n)),$$

and $H^1(A, GL(m))$ classifies projective A-modules of rank m. The map $G_m \to GL(m)$ just sends λ to λI, so on the H^1 level it sends an invertible module M to $\bigoplus_1^m M = M \oplus \cdots \oplus M$. The map to $GL(n)$ sends λ to $\lambda^{-1} I$, so it sends M to $\bigoplus_1^n M^*$, where M^* is the dual module to M. (For all this see, e.g., [27, Chap. 18].) We have $(\bigoplus_1^n M)^* = \bigoplus_1^n (M^*)$, so the images of

M are free (i.e., trivial in H^1) iff $\bigoplus_1^m M$ and $\bigoplus_1^n M$ are free. Finally, we observed in the proof of (3.4) that when $m=n$ the action of $\mathbb{Z}/2\mathbb{Z}$ on G^0 comes from the interchange action $(P, Q) \mapsto (Q, P)$ on $GL(n) \times GL(n)$; on the subgroup $(\lambda I, \lambda^{-1} I)$ this is inversion, and hence it induces inversion on $\operatorname{Pic} A = H^1(G_m)$. Thus we have computed $G(A)$:

THEOREM 3.5. *The maps $X \mapsto PXQ^{\mathrm{tr}}$ for P in $GL(m, A)$ and Q in $GL(n, A)$ form a normal subgroup of $G(A)$. When $m \neq n$, the quotient by that normal subgroup is the subgroup of $\operatorname{Pic}(A)$ consisting of those M with $\bigoplus_1^m M$ and $\bigoplus_1^n M$ free. When $m=n$, the quotient is the semi-direct product of that Picard subgroup with $(\mathbb{Z}/2\mathbb{Z})(A)$.*

The Picard group contribution to $G(A)$ will of course drop out when $\operatorname{Pic}(A)$ is trivial, as for instance when A is a local ring or a unique factorization domain. But also $\Lambda^m(\bigoplus_1^m M) = \bigotimes_1^m M$; and so, whenever $\bigoplus_1^m M$ is free, the class of M has order dividing m in $\operatorname{Pic}(A)$. Thus we really need only the torsion part of $\operatorname{Pic}(A)$ to be zero, and that is sometimes known to be true even when $\operatorname{Pic}(A)$ is nontrivial (see, e.g. [26]). In addition, we see that the Picard contribution is forced to be trivial whenever m and n are relatively prime. In any case, if T is any element of $G^0(A)$, then T^m and T^n are both of the form $X \mapsto PXQ^{\mathrm{tr}}$.

Our proof of (3.5) was an abstract cohomology computation, but one can (in several ways) use descent theory more explicitly to express elements of $G^0(A)$ as matrices. If M is an invertible A-module, it is locally free; that is, we can find finitely many elements a in A and m_a in M so that $\sum_a A = A$ and M_a is free over A_a with basis m_a. Then $B = \prod A_a$ is a faithfully flat extension of A, and $M \otimes_A B \cong \prod (M \otimes_A A_a)$ is free with generator $f = \langle m_a \rangle$. Then $B \otimes_A B \simeq \prod A_a \otimes A_b = \prod A_{ab}$, and $M \otimes_A B \otimes_A B \simeq M \otimes_A \prod A_{ab}$ has two different basis elements induced by the two maps $B \to B \otimes_A B$; in the $M \otimes A_{ab}$ component one basis element is m_a, and the other is m_b. Thus if we define c_{ab} in A_{ab}^* by the condition $m_a = c_{ab} m_b$, then $\langle c_{ab} \rangle$ in $\prod A_{ab}^* = G_m(B \otimes B)$ is the cocycle determining M (in the descent from B to A).

THEOREM 3.6. *Let M be an invertible A-module, and assume $\bigoplus_1^m M$ and $\bigoplus_1^n M$ are free. Let $\langle c_{ab} \rangle$ be defined as above. Then one can construct an invertible $m \times m$ matrix P over $B = \prod A_a$ such that in each entry the a-component is c_{ab} times the b-component. Similarly one can construct Q in $GL(n, B)$ with each a-component equal to c_{ab}^{-1} times the b-component. For such P and Q, the map $X \mapsto PXQ^{\mathrm{tr}}$ gives an element of $G^0(A)$ mapping to the Picard class of M.*

Proof. First, we construct P. By assumption we can find a basis $e_1, \ldots, e_m$ of $\bigoplus_1^m M$. We also have the basis element $f = \langle m_a \rangle$ of $M \otimes B$, and so the elements $f_j = (0, \ldots, 0, f, 0, \ldots, 0)$ form a basis of $\bigoplus_1^m M \otimes B$. We take P to be the transition matrix; that is, P_{ij} is the coefficient of $e_i \otimes 1$ when we write out f_j in the e-basis. By definition P is invertible. Inside f the component m_a is c_{ab} times m_b, so the a-component of each P_{ij} is c_{ab} times the b-component. A similar construction gives us Q, the only difference being that we take the inverse of the transition matrix.

Now we have two maps $B \to B \otimes B$, and the two images of P in $GL(m, B \otimes B)$ differ by the scalar factor $\langle c_{ab} \rangle$. Similarly the two images of Q differ by the scalar $\langle c_{ab} \rangle^{-1}$. Consider now the diagram

$$\begin{tikzcd}
1 \arrow[r] & G_m(A) \arrow[r] \arrow[d] & GL(m, A) \times GL(n, A) \arrow[r] \arrow[d] & G^0(A) \arrow[d] \\
1 \arrow[r] & G_m(B) \arrow[r] \arrow[d, shift left] \arrow[d, shift right] & GL(m, B) \times GL(n, B) \arrow[r] \arrow[d, shift left] \arrow[d, shift right] & G^0(B) \arrow[d, shift left] \arrow[d, shift right] \\
1 \arrow[r] & G_m(B \otimes B) \arrow[r] & GL(m, B \otimes B) \times GL(n, B \otimes B) \arrow[r] & G^0(B \otimes B).
\end{tikzcd}$$

The cocycle $\langle c_{ab} \rangle$ in $G_m(B \otimes B)$ maps to $(\langle c_{ab} \rangle I, \langle c_{ab} \rangle^{-1} I)$, which is precisely the matrix by which the two images of (P, Q) differ. The construction of the cohomology exact sequence [27, p. 140] then says precisely that the image of (P, Q) in $G^0(B)$ actually lies in $G^0(A)$ and maps to the class in $H^1(G_m)$ corresponding to $\langle c_{ab} \rangle$. ∎

I suppose that in some versions of the cohomology this $P \times Q^{tr}$ might correspond not to the class of M but to its inverse.

To round out these abstract computations, here is an explicit example. Let $\tau = 1 + \sqrt{-5}$, and let $A = \mathbb{Z}[\tau]$. Let M be the nonprincipal ideal $2A + \tau A$. Inverting 2 and 3, we find $M_2 = 2A_2$ and $M_3 = \tau A_3$. Thus we can take a and b to be 2 and 3, with $m_2 = 2$ and $m_3 = \tau$ (so $c_{23} = 2/\tau$). We can study 2×2 matrices, because $M \oplus M$ is free with basis

$$e_1 = (-2) \oplus \tau, \qquad e_2 = (2 - \tau) \oplus (-2).$$

Next we form $B = A_2 \times A_3$, writing its elements as pairs $\langle r, s \rangle$. Our basis element f in $M \otimes B = M_2 \times M_3$ is $\langle 2, \tau \rangle$, and we compute

$$\langle 2, \tau \rangle \oplus 0 = \langle 2, \tau \rangle e_1 + \langle \tau, \tau - 3 \rangle e_2,$$
$$0 \oplus \langle 2, \tau \rangle = \langle 2 - \tau, 3 \rangle e_1 + \langle 2, \tau \rangle e_2.$$

From this we read off the matrix

$$P = \begin{pmatrix} \langle 2, \tau \rangle & \langle 2 - \tau, 3 \rangle \\ \langle \tau, \tau - 3 \rangle & \langle 2, \tau \rangle \end{pmatrix}.$$

One can observe that the a and b components in each entry do indeed differ by the factor $2/\tau = c_{23}$. Since $m = n$, we can choose Q just to be P^{-1}. Though P has entries in B, we find of course that $PX(P^{-1})^{\mathrm{tr}}$ has coefficients in the subring A; specifically, it gives

$$\begin{aligned}
T(X_{11}) &= -2X_{11} - (\tau - 2)\, X_{12} + (\tau - 2)\, X_{21} - (\tau + 1)\, X_{22},\\
T(X_{12}) &= \tau X_{11} - 2X_{12} + 3X_{21} + (\tau - 2)\, X_{22},\\
T(X_{21}) &= -\tau X_{11} + 3X_{12} - 2X_{21} - (\tau - 2)\, X_{22},\\
T(X_{22}) &= (\tau - 3)\, X_{11} - \tau X_{12} + \tau X_{21} - 2X_{22}.
\end{aligned}$$

This T in $G^0(A)$ then maps to the class of M. In particular it cannot be written in the form PXQ^{tr} for P and Q in $GL(2, A)$, though of course it is realized in this form over B.

Our choice of B in (3.6) is not the only possible one; an element of $G^0(A)$ can be written as PXQ^{tr} over any extension B' of A that makes $M \otimes B'$ free (cf. [28, p. 348]). When A is a domain, for instance, one can choose B' to be the fraction field; this is the case treated by James [10].

4. Automorphisms of the Determinant

Any map T that preserves the determinant polynomial $\det(X_{ij})$ of course preserves the ideal generated by that polynomial, and hence it is among the maps described in Section 3. The transposition operations we know preserve the determinant. A map of the form $X \mapsto PXQ^{\mathrm{tr}}$ multiplies $\det(X)$ by the scalar $\det(PQ)$, and the homomorphism $GL(n) \times GL(n) \to G_m$ sending (P, Q) to $\det(PQ)$ factors through the quotient G^0. We observed in the example following (3.6) that we can take $Q = P^{-1}$ in (3.6) when $m = n$, and thus the Picard classes in (3.5) all have representatives preserving the determinant. From (3.4) we get at once the formal result:

THEOREM 4.1. *Let G_1 be the kernel of the homomorphism $G^0 \to G_m$ induced by $(P, Q) \mapsto \det(PQ)$. Then $G_1 \rtimes (\mathbb{Z}/2\mathbb{Z})$ is the group scheme preserving* $\det(X_{ij})$.

From (3.5) we get the explicit values:

THEOREM 4.2. *The maps $X \mapsto PXQ^{\mathrm{tr}}$ for P and Q in $GL(n, A)$ and* $\det(PQ) = 1$ *form a normal subgroup of the maps preserving* $\det(X)$. *The quotient is a semidirect product of $(\mathbb{Z}/2\mathbb{Z})(A)$ and the subgroup of* $\mathrm{Pic}(A)$ *consisting of those M with $\bigoplus_1^n M$ free.*

One historical remark is appropriate here. The computation of $\mathrm{Lie}(H_2)$ in (3.2) is new, but the Lie algebra preserving $\det(X)$ was actually computed earlier by Jacobson [8, pp. 256–258]. He did this as a Lie algebra analogue to the computation of the group itself over fields. In the absence of the group scheme context that encompasses both, however, he could not (as we did) use the Lie algebra computation as the major tool for determining the group.

5. Automorphisms of Pfaffian Ideals

We now turn from arbitrary determinants to symmetric and skew determinants. In this section we take $n(n-1)/2$ indeterminates X_{ij} with $1 \leqslant i < j < n$ and set $X_{ii} = 0$ and $X_{ji} = -X_{ij}$ to get an alternating matrix $X = (X_{ij})$. Since determinants of alternating matrices are formal squares, the natural objects of study are their square roots, the Pfaffians. For a subset I in $\{1,..., n\}$ with even cardinality $2r$, we let $\mathrm{Pf}_I(X)$ be the Pfaffian of the alternating submatrix with row and column indices in I. For $2 \leqslant r \leqslant [n/2]$, we let $H_{2r}^{-}(A)$ be the group of invertible linear maps over A preserving the ideal of these $2r \times 2r$ sub-Pfaffians. These different sub-Pfaffians (for fixed r) are independent over any field, since we can write down a matrix making any prescribed one equal 1 while all the others vanish. Hence (1.4.6) can be applied:

LEMMA 5.1. *Each H_{2r}^{-} is an affine group scheme.*

Our goal now is to determine H_{2r}^{-}, and we will follow exactly the same pattern of argument as in Section 3.

LEMMA 5.2. *One has $H_4^{-} \supseteq H_6^{-} \supseteq H_8^{-} \supseteq \cdots$.*

Proof. Suppose $2 < r \leqslant [n/2]$ and T in $H_{2r}^{-}(A)$. Write $Y_{ij} = T(X_{ij})$. If $i < j$ are two elements in a subset I of cardinality $2r$, then we have

$$\partial \mathrm{Pf}_I(X)/\partial X_{ij} = (-1)^{i+j}\, \mathrm{Pf}_{I-\{i,j\}}(X).$$

As T is in $H_{2r}^{-}(A)$, we know by (1.4.2) that the $2r \times 2r$ sub-Pfaffians of $Y = (Y_{ij})$ are A-linear combinations of those of X; say $\mathrm{Pf}_I(Y) = \sum s_I^J\, \mathrm{Pf}_J(X)$. Then

$$\begin{aligned} \pm \mathrm{Pf}_{I-\{i,j\}}(Y) &= \partial \mathrm{Pf}_I(Y)/\partial Y_{ij} \\ &= \sum \sum s_I^J(\partial \mathrm{Pf}_J(X)/\partial X_{pq})(\partial X_{pq}/\partial Y_{ij}). \end{aligned}$$

The $\partial X_{pq}/\partial Y_{ij}$ are just the coefficients of T^{-1}. Every subset of cardinality $2r-2$ can be expressed in the form $I-\{i,j\}$, and hence the $(2r-2)\times(2r-2)$ sub-Pfaffians of Y are linear combinations of those for X. Hence T is in $H^{-}_{2r-2}(A)$. ∎

LEMMA 5.3. *Let L be a field. Then the Lie algebra of H_4^- over L has dimension at most n^2.*

Proof. For $s<t<u<v$ the 4×4 Pfaffian is

$$X_{st}X_{uv}-X_{su}X_{tv}+X_{sv}X_{tu}. \qquad (*)$$

Look now at $T(X)=X+\varepsilon M(X)$, with $M(X)=(M(X)_{ij})=(\sum_{u<v} M_{ij}^{uv}X_{uv})$. Applying T to the expression $(*)$ we get back $(*)$ plus ε times the expression

$$M_{st}X_{uv}+M_{uv}X_{st}-M_{su}X_{tv}-M_{tv}X_{su}+M_{sv}X_{tu}+M_{tu}X_{sv}. \qquad (**)$$

The Lie algebra condition on the constants M_{ij}^{uv} is that the expression $(**)$ should be in the ideal generated by the Pfaffians $(*)$. As everything is homogenous of degree 2, the expressions $(**)$ then must be scalar combinations of the Pfaffians. We now just examine which terms can occur.

First, no square X_{uv}^2 ever occurs in the Pfaffians. Hence the coefficients of squares in $(**)$ must always be zero. Thus M_{ij}^{uv} is zero whenever u, v, i, and j are all distinct.

Next, we look at coefficients with one repeated index. Take, for instance, M_{ac}^{ab}. If $a\neq 1$, we look at $(**)$ for the quadruple $1, a, b, c$. No terms with repeated index occur in the Pfaffians, so the two $X_{1c}X_{ac}$ terms in $(**)$ must cancel. Thus $M_{ac}^{ab}=M_{1c}^{1b}$. Similarly M_{ab}^{bc} with $a\neq 1$ is $-M_{1a}^{1c}$, and M_{1b}^{bc} is M_{1c}^{2c}. For M_{bc}^{ab} with $a\neq 1$ we look at the $X_{1a}X_{ab}$ terms and get $M_{bc}^{ab}=-M_{1c}^{1a}$; similarly M_{bc}^{1b} with $b\neq 2$ equals M_{2c}^{12}, and $M_{2c}^{12}=M_{23}^{12}$. For M_{bc}^{ac} with $a, b\neq 1$ we get the value M_{1b}^{1a}; for M_{bc}^{1c} with $b\neq 2$ we get $-M_{2b}^{12}$, and $M_{2c}^{1c}=M_{23}^{13}$. For M_{1c}^{bc} with $b\neq 2$ we get $-M_{12}^{2b}$, and $M_{1c}^{2c}=M_{13}^{23}$. Thus all coefficients of this form can be expressed in terms of (n^2-n) of them, namely,

$$M_{1s}^{1r},\ M_{12}^{2r},\ M_{2r}^{12},\ M_{23}^{12},\ M_{23}^{13}.$$

Finally a term $X_{st}X_{uv}$ occurs in just one Pfaffian, where it has the same coefficient as $X_{sv}X_{tu}$ and $-X_{su}X_{tv}$. This gives us

$$M_{st}^{st}+M_{uv}^{uv}=M_{su}^{su}+M_{tv}^{tv}=M_{sv}^{sv}+M_{tu}^{tu}.$$

Applied to the quadruple $1, 2, a, b$, this gives us $M_{ab}^{ab}=M_{1a}^{1a}+M_{2b}^{2b}-M_{12}^{12}$; and applied to $1, 2, 3, b$, it gives us $M_{2b}^{2b}=M_{1b}^{1b}+M_{23}^{23}-M_{13}^{13}$. Thus we get n

more free terms M_{1b}^{1b} and M_{23}^{23}, and all in all we have a set of solutions of dimension at most $(n^2-n)+n=n^2$. ∎

Now for any invertible P in $GL(n, A)$ we can define $T(X)=PXP^{\mathrm{tr}}$; this T is in every H_{2r}^-. Thus we have homomorphisms $GL(n)\to H_{2r}^-$. The kernel is easily seen to be μ_2, the scalars μI with $\mu^2=1$. The quotient by this μ_2 exists [3, vol. 2, p. 15] and of course injects into H_{2r}^-. When $n=4$, the permutation T_0 in (2.3) also preserves the Pfaffian and gives us a copy of $\mathbb{Z}/2\mathbb{Z}$ inside H_4^-.

DEFINITION 5.4. For $n\neq 4$, set $G^-=GL(n)/\mu_2$. For $n=4$, set $G^-=(GL(n)/\mu_2)\rtimes\mathbb{Z}/2\mathbb{Z}$.

As μ_2 and $\mathbb{Z}/2\mathbb{Z}$ are finite group schemes, the dimension of G_L^- over every field L is still n^2, the dimension of $GL(n)$. Just as in (3.4), now, (5.3), (5.3), and (2.3) give us the information we need to apply Theorem 1.6.1, and we get our result.

THEOREM 5.5. *For every r one has $G^-\xrightarrow{\sim}H_{2r}^-$.*

As in Section 3, we can now proceed to compute the values of $G^-(A)$. For $n=4$ the factor $(\mathbb{Z}/2\mathbb{Z})(A)$ gives us maps $X_{ij}\mapsto eX_{ij}+(1-e)\,T_0(X_{ij})$ for idempotents in A; to ignore these we assume $n>4$, though what follows applies equally well to the connected component when $n=4$. The exact sequence

$$1\to\mu_2\to GL(n)\to G^-\to 1$$

gives the cohomology sequence

$$1\to\mu_2(A)\to GL(n, A)\to G^-(A)\to H^1(A,\mu_2)\to H^1(A,GL(n)).$$

Our task is to compute the final map.

First, there is an exact sequence

$$1\to\mu_2\to G_m\xrightarrow{\;2\;}G_m\to 1,$$

and the cohomology exact sequence for that gives us

$$1\to A^*/(A^*)^2\to H^1(A,\mu_2)\to \mathrm{Pic}(A)\xrightarrow{\;2\;}\mathrm{Pic}(A).$$

Thus $H^1(A,\mu_2)$ is an extension of the elements killed by 2 in $\mathrm{Pic}(A)$ by the group of unit square classes in A. This same group has occured elsewhere, for instance in Bass' work on spinor norms [1]. Furthermore, the map $\mu_2\to GL(n)$ factors through G_m, and we observed before (3.5) that the map $H^1(A, G_m)\to H^1(A, GL(n))$ sends an invertible module M to $\bigoplus_1^n M$. Thus the kernel of $H^1(A, \mu_2) \to H^1(A, GL(n))$ contains $A^*/(A^*)^2$ and also gives those invertible M with $M \otimes M \simeq A$ and $\bigoplus_1^n M$ free. We observed after (3.5) that such M will have order dividing n in $\mathrm{Pic}(A)$, so when n is odd they must be trivial. Thus we have our theorem:

THEOREM 5.6. *Assume $n > 4$. The maps $X \mapsto PXP^{\mathrm{tr}}$ for P in $GL(n, A)$ form a normal subgroup of $G^-(A)$. For odd n, the quotient by that subgroup is $A^*/(A^*)^2$. For even n, the quotient contains a copy of $A^*/(A^*)^2$, and the further quotient by that copy is the subgroup of* Pic(A) *consisting of M with* $M \otimes M$ and $\bigoplus_1^n M$ *free.*

The new feature here that was not present in (3.5) is the square class group $A^*/(A^*)^2$, which unlike $\mathrm{Pic}(A)$ can be nontrivial even for fields. Once we note these classes, we can easily recognize that they come just from multiplications of all X_{ij} by a scalar; such a scalar can be absorbed in PXP^{tr} only if it is a square. Observe, however, that we did not need to think of this when proving the basic Theorem 5.5; the cohomology computation automatically alerts us to these extra maps.

A map corresponding to an invertible module M can be constructed in the same way as in (3.6). But as we observed at that time, there are many choices for the extension B in which to realize such maps, and to illustrate that point we use a different choice here. Fix an isomorphism $M \otimes M \simeq A$. Using it as a multiplication, we can turn $B = A \oplus M$ into a (faithfully flat) A-algebra. We have $B \otimes M \xrightarrow{\sim} B$ as a B-module under the map $(a \oplus m) \otimes m' \mapsto mm' \oplus am'$. From here on we can proceed as in (3.6). That is, fix a basis element f of $B \otimes M$, and let $f_1, \ldots, f_n$ be the corresponding basis of $\bigoplus_1^n B \otimes M$. Let $e_1, \ldots, e_n$ be a basis of $\bigoplus_1^n M$, which by assumption is free. Let P be the transition matrix, so $f_j = \sum P_{ij}(1 \otimes e_i)$. The two images of P in $GL(n, B \otimes B)$ differ by a scalar factor, and that factor in $G_m(B \otimes B)$ is the cocycle defining M in the descent from B to A. Thus $X \mapsto PXP^{\mathrm{tr}}$ gives an element in $G^-(A)$ that maps to the class of M.

THEOREM 5.7. *Let n be an even number greater than* 4, *and X an $n \times n$ alternating matrix of indeterminates. The maps $X \mapsto aPXP^{\mathrm{tr}}$ for a in A and P in $GL(n, A)$ with $a^{n/2}$* $\det(P) = 1$ *form a normal subgroup of the maps preserving* Pf(X). *The quotient is the subgroup of* Pic(A) *consisting of M with $M \otimes M$ and $\bigoplus_1^n M$ free.*

Proof. Most of this follows from (5.6) and the following remarks. The one point that needs checking is that the Picard classes for maps preserving the Pfaffian ideal can all be realized by maps actually preserving $\mathrm{Pf}(X)$. We have just seen how to realize them as $X \mapsto PXP^{\mathrm{tr}}$ for P in $GL(n, B)$. As $X \mapsto PXP^{\mathrm{tr}}$ multiplies $\mathrm{Pf}(X)$ by $\det(P)$, we see that $\det P$ must be in A. It is actually invertible in A, since it is invertible in B and $A \to B$ is faithfully flat. Choose then any P' in $GL(n, A)$ with $\det(P') = (\det P)^{-1}$, and replace P by PP'. This now gives the same class in the quotient and also preserves $\mathrm{Pf}(X)$. ∎

We could also investigate the maps preserving $\det(X) = \mathrm{Pf}(X)^2$ on alternating matrices; indeed, earlier work over fields dealt with the determinant rather than the Pfaffian [18]. When A is an integral domain, of course, $\det(TX) = \det(X)$ implies $\mathrm{Pf}(TX) = \pm\mathrm{Pf}(X)$, so we get basically the same group. (The negative sign is, however, always realized, by PXP^{tr} for P with $\det P = -1$.) But if A contains an element ε with $\varepsilon^2 = 2\varepsilon = 0$, then maps preserving $\mathrm{Pf}(X)^2$ may send $\mathrm{Pf}(X)$ to many different expressions of the form $\mathrm{Pf}(X) + (\text{multiple of } \varepsilon)$. The maps preserving the determinant still form an affine group scheme, of course, but it is not so pleasant a one as we get for maps preserving $\mathrm{Pf}(X)$.

Finally, it may be worth mentioning that no noninvertible linear maps can preserve the Pfaffian [30].

6. Automorphisms of Symmetric Determinants

Following the pattern of Section 5, we now take $n(n+1)$ indeterminates X_{ij} with $1 \leqslant i \leqslant j \leqslant n$ and set $X_{ji} = X_{ij}$ to get a symmetric matrix $X = (X_{ij})$. For $2 \leqslant r \leqslant n$, let $H_r^+(A)$ be the group of invertible linear maps over A preserving the ideal of $r \times r$ minor subdeterminants of X. (As we shall see in the Appendix, we definitely want all the $r \times r$ minors, not just the principal ones.) The symmetry condition produces linear relations between these minors, for example

$$\begin{vmatrix} X_{13} & X_{23} \\ X_{14} & X_{24} \end{vmatrix} - \begin{vmatrix} X_{12} & X_{32} \\ X_{14} & X_{34} \end{vmatrix} + \begin{vmatrix} X_{12} & X_{42} \\ X_{13} & X_{43} \end{vmatrix} = 0.$$

It was proved by Runge [25], however, that for each r one can select a set of minors that span the others over $\mathbb{Z}$ and remain independent over every field. Thus (1.4.6) still can be applied:

Lemma 6.1. *Each H_r^+ is an affine group scheme.*

Here again it is usually reasonable to treat all H_r^+ together, but there are certain exceptions when 2 is a zero-divisor and r is large (see Sect. 7). The lemma we can get is this:

Lemma 6.2. *One has $H_r^+(A) \supseteq H_{r+1}^+(A)$ provided that either $r < n/2$ or 2 is not a zero divisor in A.*

Proof. As in (5.2), it is enough to take T in $H^+_{r+1}(A)$ and show that the $r\times r$ minors of $T(X)$ are linear combinations of those of X. The result of Runge, and the proof of (1.4.6), show that the span of the $r\times r$ minors of X is a direct summand in the module of all homogeneous polynomials of degree r; hence the minors of $T(X)$ will be in that span if they are so after we invert elements of A that are not zero-divisors. Thus for $r\geqslant n/2$ we may assume that 2 is actually invertible in A.

The proof will run exactly as in (3.1) and (5.2) once we show that the $r\times r$ minors of X are linear combinations of the partial derivatives of the $(r+1)\times(r+1)$ minors. This is now not quite obvious, though of course it is still clear that these partial derivatives are in the span of the $r\times r$ minors. The simplest case is when $r<n/2$, or more generally when K and L are r-element index sets with $K\cup L\neq\{1,\dots,n\}$. In this case there is an index i in neither K nor L, and the minor X_{KL} is just $\pm\partial X_{K\cup\{i\},L\cup\{i\}}/\partial X_{ii}$.

The argument is just slightly harder when $r=n-1$ and $K\cup L=\{1,\dots,n\}$. In this case there is a unique index i in $L-K$ and a unique j in $K-L$, and we get

$$\begin{aligned}\partial(\det X)/\partial X_{ij}&=(-1)^{i+j}X_{KL}+(-1)^{j+i}X_{LK}\\&=(-1)^{i+j}2X_{KL}.\end{aligned}$$

As we are now assuming 2 invertible in A, we can solve for X_{KL}.

Finally we consider the case when $r<n-1$ and $K\cup L=\{1,\dots,n\}$. We can choose then p and q in $L-K$ and s in $K-L$. Set

$$\begin{aligned}
&A=K\cup\{p\}, && B=K\cup\{q\},\\
&M=K\cup\{p\}-\{s\}, && N=K\cup\{q\}-\{s\},\\
&C=K\cup\{p,q\}-\{s\}, && D=L\cup\{s\},\\
&P=L\cup\{s\}-\{p\}, && Q=L\cup\{s\}-\{q\}.
\end{aligned}$$

The general formula, when i and j are distinct indices in $I\cap J$, is that

$$\partial X_{IJ}/\partial X_{ij}=\pm[X_{I-\{i\},J-\{i\}}+(-1)^{m(I,i,j)+m(J,i,j)}X_{I-\{j\},J-\{i\}}],$$

where $m(I,i,j)$ denotes the number of indices in I between i and j. Hence we have

$$\begin{aligned}
\partial X_{AD}/\partial X_{ps}&=\pm[X_{KL}+(-1)^{m(A,p,s)+m(D,p,s)}X_{MP}],\\
\partial X_{BD}/\partial X_{qs}&=\pm[X_{KL}+(-1)^{m(B,q,s)+m(D,q,s)}X_{NQ}],\\
\partial X_{CD}/\partial X_{pq}&=\pm[X_{NQ}+(-1)^{m(C,p,q)+m(D,p,q)}X_{MP}].
\end{aligned}$$

Distinguishing the cases where s is or is not between p and q, it is straightforward to check that the sum

$$m(A, p, s)+m(D, p, s)+m(B, q, s)+m(D, q, s)+m(C, p, q)+m(D, p, q)$$

is always even. Hence the three partials with appropriate signs will add up to exactly $2X_{KL}$. ∎

LEMMA 6.3. *Let L be a field. Then the Lie algebra of H_2^+ over L has dimension at most n^2.*

Proof. As in (3.2), the condition on $T(X)=X+\varepsilon M(X)$ is that

$$M_{ik}X_{jl}+M_{jl}X_{ik}-M_{il}X_{jk}-M_{jk}X_{il} \qquad (*)$$

be a combination of the 2×2 minors of X. Now however we have $X_{ij}=X_{ji}$ and $M_{ij}=M_{ji}$. As we promised in (3.2), we use yet a third method to investigate the conditions this time.

Let U_i be new variables, and set $X_{ij}=U_iU_j$. This gives a symmetric matrix with all 2×2 minors zero, and hence the expressions $(*)$ must vanish as functions of the U_i. It will be enough to look at $(*)$ with $i=k$. We must have then

$$\begin{aligned}\sum_{p\leqslant q} M_{ii}^{pq}\, U_pU_qU_jU_l+M_{jl}^{pq}\, U_pU_qU_i^2 \\ = M_{il}^{pq}\, U_pU_qU_jU_i+M_{ij}^{pq}\, U_pU_qU_lU_i\end{aligned}$$

for each j and l distinct from i. Only one U_i^4 term occurs, so we must have $M_{jl}^{ii}=0$. The terms involving U_i^3 give

$$\sum_{q\neq i} M_{jl}^{iq}\, U_qU_i^3=M_{il}^{ii}\, U_jU_i^3+M_{ij}^{ii}\, U_lU_i^3,$$

and thus each M_{jl}^{iq} either vanishes or equals one of the $n(n-1)$ coefficients M_{is}^{ii} with $s\neq i$. There remain only coefficients of the form M_{jl}^{jl}. For j and l not equal to 1 we take $i=k=1$ and look at the terms involving $U_jU_lU_1^2$; they give us

$$M_{11}^{11}+M_{jl}^{jl}=M_{1l}^{1l}+M_{1j}^{1j}.$$

Thus the M_{jl}^{jl} can be expressed in terms of M_{1s}^{1s} for $s=1,\ldots, n$. ∎

DEFINITION 6.4. Set $G^+=GL(n)/\mu_2$.

Just as in (5.4), a matrix P in $GL(n, A)$ defines $T(X)=PXP^{\mathrm{tr}}$ in all $M_r^+(A)$, and the map $GL(n)\to H_r^+$ has kernel μ_2, so we get injections

$G^+ \to H_r^+$. From (6.2), (6.3), and (2.2) we get the information needed to apply Theorem 1.6.1.

THEOREM 6.5. *For $r \leqslant n/2$ one has $G^+ \simeq H_r^+$. When 2 is not a zero-divisor in A, then $G^+(A) = H_r^+(A)$ for larger r as well.*

COROLLARY 6.6. *As group schemes over $\mathbb{Z}[1/2]$, one has $G^+ \simeq H_r^+$ for all r.*

The explicit evaluation of $G^+(A)$ is of course just the same as for the essentially identical group $G^-(A)$ in (5.6). Thus we can read off our final result.

THEOREM 6.7. *Let X be an $n \times n$ symmetric matrix of indeterminates. Assume 2 is not a zero-divisor in the base ring A. If n is odd, the maps $X \mapsto PXP^{\mathrm{tr}}$ with P in $GL(n, A)$ and $(\det P)^2 = 1$ are the only invertible linear maps preserving $\det X$. If n is even, the maps $X \mapsto aPXP^{\mathrm{tr}}$ with $a^n(\det P)^2 = 1$ are a normal subgroup; the quotient is the subgroup of $\mathrm{Pic}(A)$ consisting of M with $M \otimes M$ and $\bigoplus_1^n M$ free.*

Proof. The only new point to observe is that when $a^n(\det P)^2 = 1$ and $n = 2m + 1$, then $a = (a^m \det P)^{-2}$; thus if we replace P by $(a^m \det P)^{-1} P$, we can absorb the scalar. ∎

7. A DETERMINANT WITH NONSMOOTH AUTOMORPHISM GROUP

Unlike the results in Section 3 and 5, Theorem 6.5 contains restrictions on r. They stem from corresponding restrictions in Lemma 6.2. In this final section we compute an example to show that some such restrictions are indeed necessary. It will also display a type of group scheme more complicated than those we have seen before. The example we take is the smallest possible one, the group H_3^+ preserving the ideal $(\det X)$ for 3×3 symmetric X in characteristic 2.

Throughout this section we consider only rings in which $2 = 0$. The formulas are simpler if we distinguish diagonal and off-diagonal entries, so we change notation and study the matrix

$$\begin{pmatrix} X_1 & Y_3 & Y_2 \\ Y_3 & X_2 & Y_1 \\ Y_2 & Y_1 & X_3 \end{pmatrix}.$$

Since $2=0$, the determinant is just

$$D = X_1 X_2 X_3 + X_1 Y_1^2 + X_2 Y_2^2 + X_3 Y_3^2.$$

We write our linear map T as

$$U_i = TX_i = \sum a_{ij} X_j + \sum b_{ij} Y_j,$$

$$V_i = TY_i = \sum c_{ij} X_j + \sum d_{ij} \ Y_j,$$

and H_3^+ consists of those invertible T for which

$$U_1 U_2 U_3 + U_1 V_1^2 + U_2 V_2^2 + U_3 V_3^2 = \alpha D$$

for some α in A^*.

We can still get a preliminary understanding of H_3^+ by computing its Lie algebra. If we set $a_{ij} = \delta_{ij} + \varepsilon A_{ij}, \ldots, d_{ij} = \delta_{ij} + \varepsilon D_{ij}$, with $\varepsilon^2 = 0$, it is trivial to work out that the condition on T becomes

$$B_{ij} = 0,$$

$$A_{ij} = 0 \quad \text{for } i \neq j \qquad \text{and} \qquad A_{ii} = A_{jj}$$

$$C_{ij} \text{ and } D_{ij} \text{ unrestricted.}$$

In particular, $\mathrm{Lie}(H_3^+)$ in characteristic 2 has dimension 19. This shows us at once that H_3^+ is not the same as $H_2^+ = G^+$, since we know $\mathrm{Lie}(G^+)$ has dimension 9. It may be worth repeating that this is not just an indirect test distinguishing the two group schemes; rather, it is an explicit computation of some maps defined over $A = L[\varepsilon]$ that lie in $H_3^+(A)$ but not in $G^+(A)$.

For a full determination of H_3^+ in characteristic 2, we can follow the idea first used by Frobenius [6] to isolate some of the coefficients. Take (say) Y_1 and change it to $Y_1 + Z$, where Z is a new variable. This changes the determinant D to $D + X_1 Z^2$. If T is in $H_3^+(A)$, the formal identity giving αD in terms of U_i and V_i remains valid and gives

$$\begin{aligned}(U_1 + b_{11} Z)(U_2 + b_{21} Z)(U_3 + b_{31} Z) + (U_1 + b_{11} Z)(V_1^2 + d_{11}^2 Z^2)\\ + (U_2 + b_{21} Z)(V_2^2 + d_{21}^2 Z^2) + (U_3 + b_{31} Z)(V_3^2 + d_{31}^2 Z^2)\\ = \alpha(D + X_1 Z^2).\end{aligned}$$

Look first at the degree one terms in Z; they give $(b_{11} U_2 U_3 + b_{21} U_1 U_3 + b_{31} U_1 U_2 + b_{11} V_1^2 + b_{21} V_2^2 + b_{31} V_3^2) Z = 0$. Since T is invertible, the monomials in the U_i and V_i are independent, and so all the coefficients here must be zero. The same construction can be applied to Y_2 and Y_3, and so we conclude

(i) all b_{ij} are zero.

Because of this, we can drop all b_{ij} terms in our identity. This eliminates all Z^3 terms, but the Z^2 terms remain and give us

$$(d_{11}^2\, U_1 + d_{21}^2\, U_2 + d_{31}^2\, U_3)\, Z^2 = \alpha X_1 Z^2.$$

Doing this also for Y_2 and Y_3, and recalling that $U_i = \sum a_{ij} X_j$, we conclude that

(ii) $(d_{ij}^2)(a_{ij})^{\mathrm{tr}} = \alpha I$.

Now take X_1 instead and change it to $X_1 + Z$, getting

$$\begin{aligned}(U_1 + a_{11} Z)(U_2 + a_{21} Z)(U_3 + a_{31} Z) &+ (U_1 + a_{11} Z)(V_1^2 + c_{11}^2 Z^2)\\ &+ (U_2 + a_{21} Z)(V_2^2 + c_{21}^2\, Z^2) + (U_3 + a_{31} Z)(V_3^2 + c_{31}^2\, Z^2)\\ = \alpha(D + X_2 X_3 Z + Y_1^2\, Z).&\end{aligned}$$

The Z^2 terms give us

$$a_{11} a_{21} U_3 + a_{11} a_{31} U_2 + a_{21} a_{31} U_1 + c_{11}^2\, U_1 + c_{21}^2\, U_2 + c_{31}^2\, U_3 = 0.$$

Doing the same operation for X_2 and X_3, we get

(iii) $c_{ij}^2 = a_{rj} a_{sj}$, where r and s are the two indices distinct from i.

Finally, look back at the original identity for αD and consider terms involving $X_1 X_2 X_3$. Clearly they can come only from $U_1 U_2 U_3$. Since $2 = 0$, signs are irrelevant, and we get exactly $\det(a_{ij})\, X_1 X_2 X_3$. Hence

(iv) $\det(a_{ij}) = \alpha$.

Conversely, now, it is a straightforward computation to verify that conditions (i), (ii), (iii), (iv) do indeed imply $\alpha D = U_1 U_2 U_3 + U_1 V_1^2 + U_2 V_2^2 + U_3 V_3^2$.

THEOREM 7.1. *Consider 3×3 symmetric matrices over a ring A where $2 = 0$. The elements in $H_3^+(A)$ are those of the form*

$$TX_i = \sum a_{ij} X_j,$$

$$TY_i = \sum c_{ij} X_j + \sum d_{ij} Y_j,$$

where (a_{ij}) is an invertible matrix, $(d_{ij}^2)(a_{ij})^{\mathrm{tr}} = \det(a_{ij})\, I_3$, and $c_{ij}^2 = a_{rj} a_{sj}$ where r and s are the two indices distinct from i.

We still need to understand better how H_3^+ here is related to G^+. As H_3^+ for characteristic 2 is defined over the perfect field $\mathbf{F}_2$, one knows abstractly

[27, p. 53] that it contains a smooth subgroup $(H_3^+)_{\text{red}}$ of its same dimension, and in fact that is just what we want:

THEOREM 7.2. *For* 3×3 *symmetric matrices in characteristic* 2, *one has* $G^+ \simeq (H_3^+)_{\text{red}}$.

Proof. The map sending P to $(X \mapsto PXP^{\text{tr}})$ of course still yields an injection of G^+ into H_2^+. As G^+ is smooth, the image lands in $(H_3^+)_{\text{red}}$. Now in (7.1) we get a homomorphism $H_3^+ \to GL(3)$ by sending T to (a_{ij}). Computation shows that the composite

$$GL(3) \to G^+ \to H_3^+ \to GL(3)$$

sends (P_{ij}) to (P_{ij}^2). Hence $H_3^+ \to GL(3)$ is a group scheme epimorphism. Its kernel consists of those T with $a_{ij} = \delta_{ij}$; they have $c_{ij}^2 = 0$ and $d_{ij}^2 = \delta_{ij}$, and thus they form just a finite connected group scheme of height 1. It follows that H_3^+ is connected and has the same dimension as $GL(3)$. Automatically $(H_3^+)_{\text{red}}$ inherits these properties. As G^+ is a smooth subgroup of the same dimension, $G^+ = (H_3^+)_{\text{red}}$. ∎

In a field, or more generally in a ring without nilpotents, any set of values satisfying the equations defining H_3^+ will also satisfy those defining $(H_3^+)_{\text{red}}$. Hence:

COROLLARY 7.3. *If A is a ring where $2 = 0$ but A has no nilpotents, then* $H_3^+(A) = G^+(A)$.

It would be nice to know whether such statements remain true for lager matrices.[1]

APPENDIX

There are several different ways to formulate "preserver" questions, and this Appendix will discuss how they are related to the formulation we have adopted. To begin with a minor point, we are studying only the *invertible* matrices preserving polynomials or ideals. In certain degenerate cases there may also be noninvertible maps preserving them. This possible degeneracy is best studied separately; indeed, it can often be determined immediately once the automorphisms are known, as I have shown in [30]. Also, the primary application of preserver information is in descent theory, for which only the automorphisms matter. But the whole question is in any case irrelevant for our particular topic, because the determinant in fact is not preserved by any noninvertible maps [30].

[1] *Note added in proof.* See now *Proc. Amer. Math. Soc.* **93**, 583–589.

A more significant distinction arises from the different senses that may be given to the phrase "preserving the polynomial f." We have adopted the usual algebraist's definition in which a polynomial is a formal expression. Thus our requirement that $f \circ T = f$ means that the coefficients on both sides should agree. It is possible instead to interpret f as a function and ask that its values be preserved. This in turn can be done in two ways. First, one can take f to define a function on all extensions of k (as it does) and ask that all such values be preserved. It is easy to see that this is equivalent to formal preservation of the polynomial. Alternately, however, one might ask that the map preserve the values of f just on n-tuples from the base ring k itself. If k is an infinite field, of course, it is well known that this still implies that f is formally preserved. It is also true that for the determinant, preservation of values over k always implies formal preservation; this can be seen by comparing our results with the simultaneous, independent work of McDonald [19], who has studied the maps preserving k-values of the determinant. Still, it is possible to find cases where these two formulations give different answers. We can, for instance, take $f(X, Y) = XY^p - XY$ over $k = \mathbb{Z}/p\mathbb{Z}$. The only maps preserving f send X to aX and Y to $a^{-1}Y$ with $a^{p-1} = 1$. But f vanishes identically for X and Y in k, so all linear maps over k preserve its k-values.

As this example suggests, differences between the two formulations arise only when k by itself is too small to mark the distinction between two polynomials that are formally different; and the extra maps preserving k-values are always exceptional. More precisely, this means that they will not fit into a functor; that is, there will be a homomorphism $k \to k'$ such that the exceptional maps preserving k-values no longer preserve values over k'. Thus it is only when we ask for maps preserving f formally that we can be sure of getting an answer that is systematic and well behaved as the base ring varies.

A similar distinction arises when we deal with maps preserving an ideal I. The definition we have adopted requires that T formally preserve the vanishing of the polynomials generating I. If we set $V_I(A)$ to be the elements in A^n where all polynomials in I vanish, we could instead ask for the T sending $V_I(A)$ to $V_I(A)$. Again it is easy to see that if this holds for all extensions of k, then T must formally preserve I. But there may be other maps sending the single set $V_I(k)$ to itself, though as before they are necessarily exceptional and will not fit into a functor. Here, though, such exceptional maps may occur even over infinite fields, because the equations defining I simply may not have many solutions in k. Suppose, for instance, that over $k = \mathbb{R}$ we take I to be generated by the quaternion reduced norm $X_1^2 + X_2^2 + X_3^2 + X_4^2$. The maps that formally preserve I are those in the general orthogonal group. But of course every linear map preserves the origin, which is the only point in $V_I(\mathbb{R})$.

A more subtle source of exceptional values is the presence of nilpotent elements in our rings. An alternating matrix with zero determinant, for instance, may not have zero Pfaffian. That is, the condition $f=0$ is formally different from $f^2=0$, though they have the same solutions over fields. More generally, there will be many different ideals I with the same $V_I(k)$. For a classical example of this we may take a 3×3 symmetric matrix of indeterminates $X=(X_{ij})$ (where $X_{ij}=X_{ji}$). Over fields k it is well known that a 3×3 symmetric matrix will have rank $\leqslant 1$ provided that its determinant and all 2×2 principal minors vanish. In other words, if I is the ideal

$$(X_{11}X_{22}-X_{12}^2, X_{11}X_{33}-X_{13}^2, X_{22}X_{33}-X_{23}^2, \det(X)),$$

then $V_I(k)$ consists of the symmetric matrices of rank $\leqslant 1$. But I does not actually contain the nonprincipal 2×2 minors; it only contains their squares. This automatically means that there is a symmetric 3×3 matrix over some k-algebra where the determinant and principal 2×2 minors vanish but the other 2×2 minors are nonzero. The congruence map $X\mapsto PXP^{\mathrm{tr}}$ for

$$P=\begin{pmatrix}1 & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}$$

preserves the set of symmetric matrices of rank $\leqslant 1$, but it does not preserve I; indeed, the $(1, 3)$-principal minor of PXP^{tr} is

$$(X_{11}X_{33}-X_{13}^2)+(X_{22}X_{33}-X_{23}^2)+2(X_{12}X_{33}-X_{13}X_{23}),$$

of which the last term is not in I.

One might still ask when (if ever) we can tell *a priori* that all k-linear changes of variable preserving $V_I(k)$ preserve I. Trivially, of course, this holds if I is *defined* to be the polynomials vanishing on a set in k^n (cf. Dixon [5]). The question is how to translate this condition into more intrinsic properties of I. In general this question has a good answer only over algebraically closed fields: there it suffices that $k[X]/I$ have no nilpotent elements. (Indeed, if $f\circ T$ vanishes on $V_I(k)$, then some $(f\circ T)^m$ is in I by the Hilbert Nullstellensatz, and so by assumption $f\circ T$ is in I.) In particular this holds if I is a prime ideal. Over other infinite fields the question is trickier, though it always comes down to whether the equations defining I have "enough" solutions in k. Formally, it is enough if $\bar{k}[X]/I\bar{k}[X]$ has no nilpotents and $V_I(k)$ is "Zariski-dense" in $V_I(\bar{k})$.

Determinant ideals satisfy these requirements. That is, let (X_{ij}) be an $m\times n$ matrix of indeterminates over a field k, and let I be the ideal generated by all $r \times r$ minors. Then I is a prime ideal [7, p. 1045]. Furthermore, we can write the matrices of rank less than r over $\bar{k}$ as

$$P\begin{pmatrix} I_{r-1} & 0 \\ 0 & 0 \end{pmatrix} Q$$

for matrices P and Q over $\bar{k}$. If k is an infinite field, we can approximate P and Q in the Zariski topology by matrices over k, and thus $V_I(k)$ here is dense in $V_I(\bar{k})$. Hence over any infinite field the changes of variable preserving this $V_I(k)$ are the same as those actually preserving I. The same result is similarly true for minors of symmetric or alternating matrices [11]. It should be emphasized, however, that these results are not actually needed for this paper.

References

1. H. Bass, Clifford algebras and spinor norms over a commutative ring, *Amer. J. Math.* **96** (1974), 156–206.
2. M. Demazure and P. Gabriel, "Groupes Algébriques I," North-Holland, Amsterdam, 1970.
3. M. Demazure, A. Grothendieck, *et al.*, "SGA 3: Schémas en Groupes," Lecture Notes in Math. Vols. 151, 152, 153, Springer-Verlag, New York, 1970.
4. J. Dieudonné, Sur une généralisation du groupe orthogonal à quatre variables, *Arch. Math.* **1** (1948–1949), 282–287.
5. J. D. Dixon, Rigid embedding of simple groups in the general linear group, *Canad. J. Math.* **29** (1977), 384–391.
6. G. Frobenius, Über die Darstellung der endlichen Gruppen durch lineare Substitutionen, Berlin Sitzungsber., 1897, pp. 994–1015; Abhandlungen III, pp. 82–103.
7. M. Hochster and J. A. Eagon, Cohen–Macaulay rings, invariant theory, and the generic perfection of determinantal loci, *Amer. J. Math.* **93** (1971), 1020–1058.
8. N. Jacobson, "Structure and Representations of Jordan Algebras," Colloq. Publ. No. 39, Amer. Math. Soc., Providence, R. I., 1968.
9. D. G. James, On the orthogonal group $O(H(A \oplus A), A)$, *J. London Math. Soc.* (2) **4** (1972), 584–586.
10. D. G. James, On the automorphisms of $\det(X_{ij})$, *Math. Chronicle* **9** (1980), 35–40.
11. H. Kleppe and D. Laksov, The algebraic structure and deformation of Pfaffian schemes, *J. Algebra* **64** (1980), 167–189.
12. M. H. Lim, Linear transformations on symmetric matrices, *Linear Multilinear Algebra* **7** (1979), 47–57.
13. M. Marcus, Linear operations on matrices, *Amer. Math. Monthly* **69** (1962), 837–847.
14. M. Marcus, Linear transformations on matrices, *J. Res. Nat. Bur. Standards B* **75** (1971), 107–113.
15. M. Marcus and J. Chollet, Linear groups defined by decomposable tensor equalities, *Linear Multilinear Algebra* **8** (1980), 207–212.
16. M. Marcus and F. May, On a theorem of. I. Schur concerning matrix transformations, *Arch. Math.* **11** (1960), 401–404.
17. M. Marcus and B. N. Moyls, Linear transformations on algebras of matrices, *Canad. J. Math.* **11** (1959), 61–66.

18. M. MARCUS AND R. WESTWICK, Linear maps on skew symmetric matrices: The invariance of elementary symmetric functions, *Pacific J. Math.* **10** (1960), 917–924.
19. B. R. MCDONALD, "R-linear Endomorphisms of $(R)_n$ Preserving Invariants," Memoirs. Amer. Math. Soc. **46** (No. 287), 1983.
20. H. MINC, Linear transformations on matrices: Rank 1 preservers and determinant preservers, *Linear Multilinear Algebra* **4** (1977), 265–272.
21. M. ORZECH AND L. RIBES, On residual finiteness and the Hopf property in rings, *J. Algebra* **15** (1970), 81–88.
22. S. PIERCE, Discriminant preserving linear maps, *Linear Multilinear Algebra* **8** (1979), 101–114.
23. J. RACKUSIN AND W. WATKINS, Linear maps on matrices: The invariance of symmetric polynomials, *Linear Algebra Appl.* **17** (1977), 269–276.
24. A. ROSENBERG AND D. ZELINSKY, Automorphisms of separable algebras, *Pacific J. Math.* **11** (1961), 1109–1117.
25. C. RUNGE, Die linearen Relationen zwischen den verschiedenen Subdeterminanten symmetrischer Systeme, *J. Reine Angew. Math.* **93** (1882), 319–327.
26. W. C. WATERHOUSE, Class groups of rings of invariants, *Proc. Amer. Math. Soc.* **67** (1977), 23–26.
27. W. C. WATERHOUSE, "Introduction to Affine Group Schemes," Graduate Texts in Math., No. 66, Springer-Verlag, New York, 1979.
28. W. C. WATERHOUSE, Automorphisms of $GL_n(R)$, *Proc. Amer. Math. Soc.* **79** (1980), 347–351.
29. W. C. WATERHOUSE, Automorphisms of quotients of $\prod GL(n_i)$, *Pacific J. Math.* **102** (1982), 221–233.
30. W. C. WATERHOUSE, Invertibility of linear maps preserving matrix invariants, *Linear Multilinear Algebra* **13** (1983), 105–113.
31. W. C. WATERHOUSE, Twisted forms of the determinant, *J. Algebra* **86** (1984), 60–75.
32. W. J. WONG, Maps on simple algebras preserving zero products. I. The associative case, *Pacific J. Math.* **89** (1980), 229–247.

Printed in Belgium